

本周周报(11.26-12.2):

解聪

本周工作:

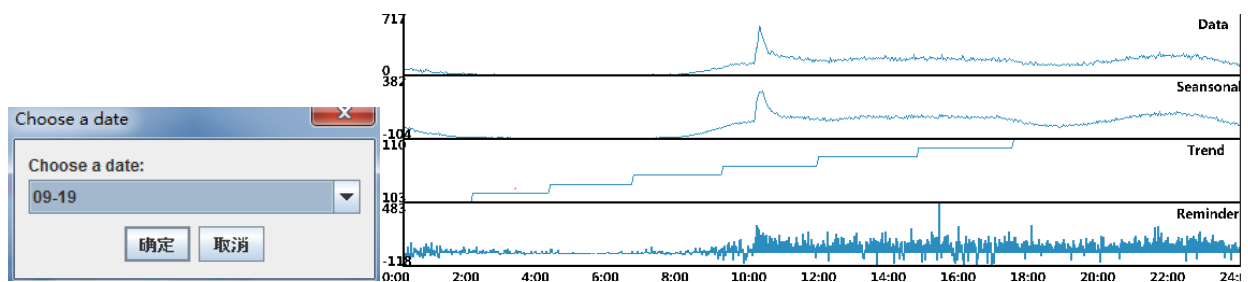
1. 时序用户行为分析

数据库接入

本周将获得的总共 2 千万条一周的交易数据导入数据库，并在数据库的基础上修改了程序读取处理数据的部分。采用文件处理的缺点在于每次读文件的时候不能随机读取，只能顺序地查找到所要读取的记录。而且数据存储多个文件中，来回切换也是程序实现非常不方便。

分析流程的调整

将 STL 的视图改为动态可变。不再显示全部一周的数据并让用户框选。系统先让用户选取某一天，再以一天为单位显示 STL 曲线。对于用户，则是以天为单位进行异常交易数据的分析。

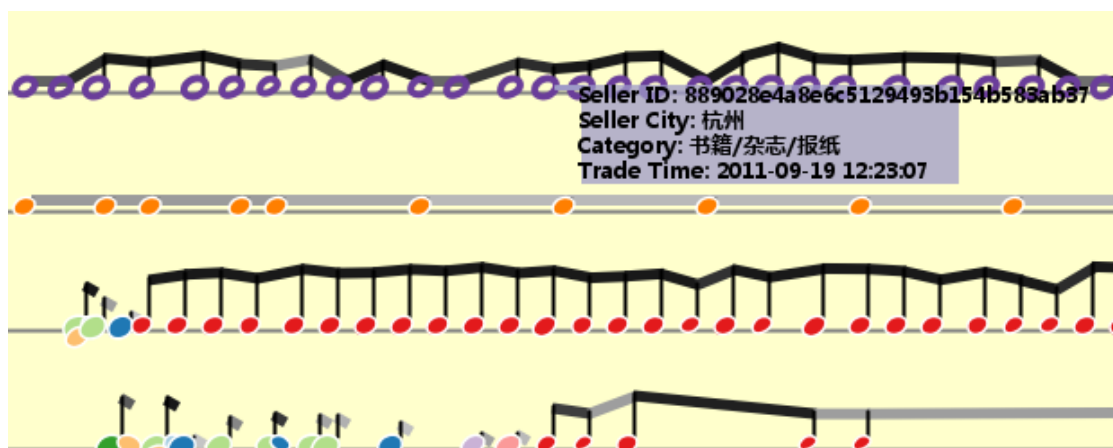


将 STL 视图改为了只显示一天的异常信息。一是使用户更清晰的观察某一天的数据。二是明确了分析流程，用户首先选择对应的日期，再对该日期进行可视分析。如果显示全部一周的曲线，某些细节信息会在全局信息被忽略掉。

可视设计的修改

本周修改了一些可视设计，使系统便于用户的理解与交互。

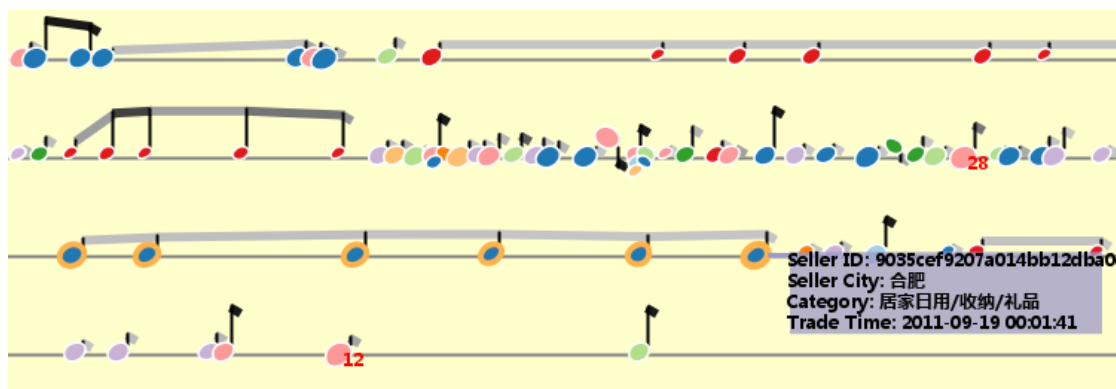
为了表示一些城市的异常信息，使用空心的符点表示。



如上图所示，其中第一行的交易的类目属于书籍类目，属于实体商品。相比于第二行的橙

色的交易，其价格也是正常范围之内波动（紫色的符尾的趋势有所起伏，而橙色交易的符尾维持在低位，表示每次的交易价格很低）。但是紫色交易的异常点在于买家地址是空，表示其没有发货信息，就像之前分析的一样，这可能就是虚假交易。使用空心的音符表示出其地点异常的信息，以吸引分析师的注意。

为了表示出相同卖家 ID 与相同买家 ID 的一对一的多次交易，采用了在用户交互时，将相应的所有音符高亮的可视化方式：



如图，用户选择了其中一个音符，而系统则使用橙色边框高亮了一系列相关的交易，表示每组交易不仅卖家 ID 相同，买家 ID 也是一样。可以发现在短时间内，相同的卖家与买家进行了多次交易，可以认为其中隐藏着一定的用户行为。因此，在用户交互时，将与当前选择的交易相关的一些交易也高亮显示有助于帮助用户找到买卖家之间的对应关系。

2. DataV 组件

本周完成雷达图的绘制与完善，并修改了一些组件的相应的数据格式使其便于用户使用。

下周工作：

1. 继续阅读论文，尝试找到一些结合高维数据与时序数据的的分析方法。
2. 完善时序用户行为分析系统中现有可视分析流程。
3. 修改浙大学报工学版的投稿论文。